

#6

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re application of: **Arei KOBAYASHI, et al.**

Serial No.: **Not Yet Assigned**

Filed: **December 26, 2001**

For: **METHOD AND SYSTEM FOR CODE PROCESSING OF DOCUMENT DATA**

CLAIM FOR PRIORITY UNDER 35 U.S.C. 119

Commissioner for Patents
Washington, D.C. 20231

December 26, 2001

Sir:

The benefit of the filing date of the following prior foreign application is hereby requested for the above-identified application, and the priority provided in 35 U.S.C. 119 is hereby claimed:

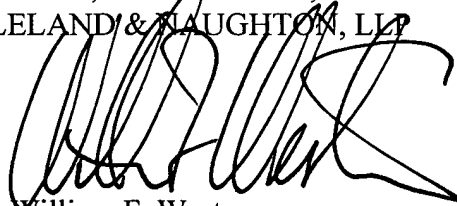
Japanese Appln. No. 2000-397002, filed December 27, 2000

In support of this claim, the requisite certified copy of said original foreign application is filed herewith.

It is requested that the file of this application be marked to indicate that the applicants have complied with the requirements of 35 U.S.C. 119 and that the Patent and Trademark Office kindly acknowledge receipt of said certified copy.

In the event that any fees are due in connection with this paper, please charge our Deposit Account No. 01-2340.

Respectfully submitted,
ARMSTRONG, WESTERMAN, HATTORI
MCLELAND & LAUGHTON, LLP



William F. Westerman
Reg. No. 29,988

Atty. Docket No.: 011736
Suite 1000, 1725 K Street, N.W.
Washington, D.C. 20006
Tel: (202) 659-2930
Fax: (202) 887-0357
WFW/ll



PATENT OFFICE
JAPANESE GOVERNMENT



This is to certify that the annexed is a true copy of the following application as filed with this Office.

Date of Application: December 27, 2000

Application Number: 397002/2000

Applicant(s): KDDI Corporation

September 10, 2001

Commissioner,
Patent Office

Kozo OIKAWA(Official Seal)

Certificate Issuance No.2001-3083355

[Document] Application for Patent
[Reference Number] P-8612
[Filing Date] December 27, 2000
[Recipient] Commissioner, Patent Office
[IPC Number] G06F 17/22
H04L 29/06

[Inventor(s)]
[Address] c/o KDD R&D Laboratories Inc.
1-15, Ohara 2-chome, Kamifukuoka-shi,
Saitama
[Name] Arei KOBAYASHI
[Inventor(s)]
[Address] c/o KDD R&D Laboratories Inc.
1-15, Ohara 2-chome, Kamifukuoka-shi,
Saitama
[Name] Satoru TAKAGI
[Applicant]
[Identification Number] 000208891
[Name] DDI Corporation
[Attorney]
[Identification Number] 100074930
[Patent Attorney]
[Name] Keiichi YAMAMOTO
[General Fee]
[Deposition Account Number] 001742
[Amount] 21000 yen
[List of Attached Document]
[Document] Specification 1
[Document] Drawings 1
[Document] Abstract 1
[General Power Number] 0016646
[Necessity of Proof] Necessary

日 本 国 特 許 庁
JAPAN PATENT OFFICE



別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office

出 願 年 月 日

Date of Application:

2000年12月27日

出 願 番 号

Application Number:

特願2000-397002

出 願 人

Applicant(s):

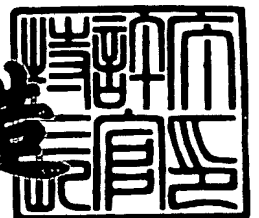
ケイディーディーアイ株式会社



2001年 9月10日

特 許 庁 長 官
Commissioner,
Japan Patent Office

及 川 耕 造



出証番号 出証特2001-3083355

【書類名】 特許願

【整理番号】 P-8612

【提出日】 平成12年12月27日

【あて先】 特許庁長官殿

【国際特許分類】 G06F 17/22
H04L 29/06

【発明者】

【住所又は居所】 埼玉県上福岡市大原二丁目1番15号 株式会社ケイデ
ィディ研究所内

【氏名】 小林 亜令

【発明者】

【住所又は居所】 埼玉県上福岡市大原二丁目1番15号 株式会社ケイデ
ィディ研究所内

【氏名】 高木 悟

【特許出願人】

【識別番号】 000208891

【氏名又は名称】 株式会社ディーディーアイ

【代理人】

【識別番号】 100074930

【弁理士】

【氏名又は名称】 山本 恵一

【手数料の表示】

【予納台帳番号】 001742

【納付金額】 21,000円

【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

【包括委任状番号】 0016646

特 2 0 0 0 - 3 9 7 0 0 2

【プルーフの要否】 要

【書類名】 明細書

【発明の名称】 文書データ符号化方法、システム及びプログラムを記録した記録媒体

【特許請求の範囲】

【請求項 1】 拡張可能なテキスト形式の記述言語で記載された文書データを符号化した符号化データを送信し、受信した該符号化データを復号化して前記文書データを復元する文書データ符号化方法であって、

前記文書データの符号化及び復号化は、拡張可能なテキスト形式の記述言語で記載された符号化変換テーブルに基づいて行われ、

前記符号化変換テーブルには、

要素の名称、該要素の型、該要素の値、該要素に付随する属性の出現回数、該属性の名称、該属性の型及び／又は該属性の値の項目を識別する接頭辞として割り当てられた符号長及び符号と、

それぞれの前記項目の内容に対して割り当てられた符号長及び符号と、

要素の親子構造関係を示す符号長及び符号と

が定義されていることを特徴とする文書データ符号化方法。

【請求項 2】 前記符号化変換テーブルには、他の複数の符号化変換テーブルのリンク情報が定義されていることを特徴とする請求項 1 に記載の方法。

【請求項 3】 拡張可能なテキスト形式の記述言語の文書データを受信する手段、該文書データを符号化する手段、及び該符号化した符号化データを送信する手段を有する符号化サーバと、前記符号化データを受信する手段、及び該符号化データを復号化する手段を有する復号化クライアントとを有する文書データ符号化システムであって、

前記文書データの符号化及び復号化は、拡張可能なテキスト形式の記述言語で記載された符号化変換テーブルに基づいて行われ、

前記符号化変換テーブルには、

要素の名称、該要素の型、該要素の値、該要素に付随する属性の出現回数、該属性の名称、該属性の型及び／又は該属性の値の項目を識別する接頭辞として割り当てられた符号長及び符号と、

それぞれの前記項目の内容に対して割り当てられた符号長及び符号と、
要素の親子構造関係を示す符号長及び符号と
が定義されていることを特徴とする文書データ符号化システム。

【請求項4】 前記符号化サーバは、前記符号化変換テーブルを他のサーバから取得する手段を有することを特徴とする請求項3に記載のシステム。

【請求項5】 前記復号化クライアントの前記符号化データを復号化する手段は、該符号化データを直接的に処理することができる文書専用処理エンジンであることを特徴とする請求項3又は4に記載のシステム。

【請求項6】 拡張可能なテキスト形式の記述言語で記載された符号化変換テーブルに基づいて、拡張可能なテキスト形式の記述言語の文書データを符号化データに符号化する文書データ符号化プログラムを記録した記録媒体であって、
前記符号化変換テーブルには、

要素の名称、該要素の型、該要素の値、該要素に付随する属性の出現回数、該属性の名称、該属性の型及び／又は該属性の値の項目を識別する接頭辞として割り当てられた符号長及び符号と、

それぞれの前記項目の内容に対して割り当てられた符号長及び符号と、
要素の親子構造関係を示す符号長及び符号と
が定義されていることを特徴とする文書データ符号化プログラムを記録した記録媒体。

【請求項7】 拡張可能なテキスト形式の記述言語で記載された符号化変換テーブルに基づいて、符号化データを、拡張可能なテキスト形式の記述言語の文書データに復号化する文書データ復号化プログラムを記録した記録媒体であって、

前記符号化変換テーブルには、
要素の名称、該要素の型、該要素の値、該要素に付随する属性の出現回数、該属性の名称、該属性の型及び／又は該属性の値の項目を識別する接頭辞として割り当てられた符号長及び符号と、

それぞれの前記項目の内容に対して割り当てられた符号長及び符号と、
要素の親子構造関係を示す符号長及び符号と

が定義されていることを特徴とする文書データ復号化プログラムを記録した記録媒体。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は、文書データ符号化方法、システム及びプログラムを記録した記録媒体に関する。

【0002】

【従来の技術】

従来、テキスト形式の記述言語の文書データ（例えばHTML: HyperText Markup Language）を伝送する際に、伝送のデータ量を削減するために、当該文書データを符号化する方法があった。このために、当該記述言語と符号化データとが1対1に対応する符号化変換テーブルを、予め、送信側装置と受信側装置とに備えておく必要がある。受信側装置は、符号化変換テーブルに基づいて、受信した符号化データを逐次復号させ、元の文書データを復元させ、その文書データに基づいて処理が行われていた。

【0003】

【発明が解決しようとする課題】

しかしながら、定型の符号化変換テーブルでは、汎用性及び拡張性が無いために、拡張可能なテキスト形式の記述言語、例えばSGML: Standard Generalized Markup Language又はXML: eXtensible Markup Languageには対応できないという課題があった。即ち、拡張された場合、その都度、符号化変換テーブルを変更しなければならない。

【0004】

そこで、本発明は、拡張可能なテキスト形式の記述言語についての符号化を可能とする方法を提供することを目的とする。

【0005】

【課題を解決するための手段】

本発明は、拡張可能なテキスト形式の記述言語で記載された文書データを符号

化した符号化データを送信し、受信した該符号化データを復号化して文書データを復元する文書データ符号化方法であって、文書データの符号化及び復号化は、拡張可能なテキスト形式の記述言語で記載された符号化変換テーブルに基づいて行われ、符号化変換テーブルには、要素の名称、該要素の型、該要素の値、該要素に付随する属性の出現回数、該属性の名称、該属性の型及び／又は該属性の値の項目を識別する接頭辞として割り当てられた符号長及び符号と、それぞれの項目の内容に対して割り当てられた符号長及び符号と、要素の親子構造関係を示す符号長及び符号とが定義されている方法である。この符号化変換テーブルにより、拡張可能なテキスト形式の記述言語についての符号化が可能となる。

【0006】

本発明の他の実施形態によれば、符号化変換テーブルには、他の複数の符号化変換テーブルのリンク情報が定義されていることも好ましい。これにより、符号化変換テーブル自体も拡張可能にすることができるので、拡張可能なテキスト形式の記述言語で記載された文書データに対応することが可能となる。

【0007】

また、本発明は、拡張可能なテキスト形式の記述言語の文書データを受信する手段、該文書データを符号化する手段、及び該符号化した符号化データを送信する手段を有する符号化サーバと、符号化データを受信する手段、及び該符号化データを復号化する手段を有する復号化クライアントとを有する文書データ符号化システムであって、文書データの符号化及び復号化は、拡張可能なテキスト形式の記述言語で記載された符号化変換テーブルに基づいて行われ、符号化変換テーブルには、要素の名称、該要素の型、該要素の値、該要素に付随する属性の出現回数、該属性の名称、該属性の型及び／又は該属性の値の項目を識別する接頭辞として割り当てられた符号長及び符号と、それぞれの項目の内容に対して割り当てられた符号長及び符号と、要素の親子構造関係を示す符号長及び符号とが定義されているシステムである。

【0008】

本発明の他の実施形態によれば、符号化サーバは、符号化変換テーブルを他のサーバから取得する手段を有するものであってもよい。

【0009】

本発明の他の実施形態によれば、復号化クライアントの符号化データを復号化する手段は、該符号化データを直接的に処理することができる文書専用処理エンジンであることが好ましい。これにより、受信した符号化データから元の文書データを復元させる必要がない。加えて、前述したように本発明による符号化データには、要素の親子関係構造も符号化されているために、復元された元の文書データに対して、更にパーシング処理を施す必要もない。これは、例えば携帯電話機のような低い処理能力しか有さない受信側装置にとって処理負荷が小さいという効果をもたらす。

【0010】

更に、本発明は、拡張可能なテキスト形式の記述言語で記載された符号化変換テーブルに基づいて、拡張可能なテキスト形式の記述言語の文書データを符号化データに符号化する文書データ符号化プログラムを記録した記録媒体であって、符号化変換テーブルには、要素の名称、該要素の型、該要素の値、該要素に付随する属性の出現回数、該属性の名称、該属性の型及び／又は該属性の値の項目を識別する接頭辞として割り当てられた符号長及び符号と、それぞれの項目の内容に対して割り当てられた符号長及び符号と、要素の親子構造関係を示す符号長及び符号とが定義されている文書データ符号化プログラムを記録した記録媒体である。

【0011】

更に、本発明は、拡張可能なテキスト形式の記述言語で記載された符号化変換テーブルに基づいて、符号化データを、拡張可能なテキスト形式の記述言語の文書データに復号化する文書データ復号化プログラムを記録した記録媒体であって、符号化変換テーブルには、要素の名称、該要素の型、該要素の値、該要素に付随する属性の出現回数、該属性の名称、該属性の型及び／又は該属性の値の項目を識別する接頭辞として割り当てられた符号長及び符号と、それぞれの項目の内容に対して割り当てられた符号長及び符号と、要素の親子構造関係を示す符号長及び符号とが定義されている文書データ復号化プログラムを記録した記録媒体である。

【0012】

【発明の実施の形態】

以下では、図面を用いて、本発明の実施形態を詳細に説明する。

【0013】

図1は、本発明の基本的な符号化方法の構成図である。エンコーダは、XML形式のプレーンテキストデータを、XML形式の符号化変換テーブルを用いて符号化し、符号化データを出力する。一方、デコーダは、その符号化データを、同じXML形式の符号化変換テーブルを用いて、最初のプレーンテキストデータを出力する。

【0014】

図2は、XML文書と、本発明による符号化変換テーブルと、該テーブルを用いて該文書を符号化した符号化データとの一例を表している。本発明による特徴ある部分を以下に説明する。

【0015】

符号化変換テーブルは、ヘッダ（インデックス）部分とボディ（本体）部分とからなる。

```
<xeus>
  <header>...</header>
  <body>...</body>
</xeus>
```

【0016】

ヘッダ部分には、要素・属性の名称、文字列及び数値等を区別するために、接頭辞の符号及び符号長を割り当てることができる。

```
<prefix bit="2">          [符号長]
  <name code="00">        [名称に対する符号]
  <number code="01"/>     [数値に対する符号]
  <char code="10"/>       [文字列に対する符号]
</prefix>
```

【0017】

ヘッダ部分には、割り当てる符号定義の進法を定義することができる。

`<xeus code-type="binary" >` [2進法で記述]

【0018】

ボディ部分には、文書の論理構造と符号化変換情報とを記述をすることができる。

【0019】

ボディ部分には、要素の名称に対して符号長及び符号を割り当てることができる。

`<element name="sample" bit="3" code="001">`

[sampleという要素に3bitの符号長で、001を割り当てる]

【0020】

ボディ部分には、要素に付随する属性に対して符号長を割り当てることができる。

`<attlist bit="2">`

[属性に2ビットを割り当てる]

【0021】

要素に付随する属性の名称に対して、型、値の型、候補を記述できる。

`<attr name="attr_sample" type="implied" code="10">`

[attr_sampleという名称の属性が存在し、型はimpliedである]

`<value>`

`<number bit="4" data="UI" qt="1" />`

`</value>`

【0022】

図3は、複数の符号化変換テーブルを結合したインデックステーブルの説明図である。本発明が対象としている記述言語は、拡張可能なテキスト形式の記述言語であるので、拡張される毎に符号化変換テーブルを作成し直すことは、非常に負担となる。従って、図6のように、複数の符号化変換テーブルを結合することにより、符号化変換テーブルを作成し直すという負担を無くすることができる。

【0023】

図 3 によれば、ヘッダ部分には、複数の変換テーブルを拡張するためのメタ情報（接頭辞符号の符号・符号長、要素の指定、名前空間の指定、符号化変換テーブルへのリンク情報）を定義することができる。

```
<namespaces>
```

```
  <namespace bit="1" name="a1" prefix="a" code="0" xlink:href="a.xeus"
```

```
/>
```

[接頭辞符号=1、拡張対象の要素=<a1>、接頭辞=a、接頭辞符号=0、
符号化変換テーブルへのリンク=a.xeus]

```
  <namespace bit="1" name="a1" prefix="b" code="1" xlink:href="b.xeus"
```

```
/>
```

```
</namespaces>
```

【 0 0 2 4 】

図 4 は、本発明による第 1 の実施形態のシステム構成図である。図 1 のエンコーダをプロキシサーバとして、図 1 のデコーダをクライアントとして実現したものである。クライアントは、エンコードサーバを介して既存のサーバへ要求する。該サーバは、XML 形式のプレーンテキストデータをエンコードサーバへ送信し、該エンコードサーバは、該データを符号化変換テーブルを用いて符号化する。その符号化されたデータは、クライアントへダウンロードされる。該クライアントは、符号化変換テーブルを用いて復号化し、XML 形式のプレーンテキストデータを生成する。そして、そのデータを処理することが可能となる。これにより、XML 形式のプレーンテキストデータを提供する既存のサーバに変更を加えることなく、プロキシサーバによって符号化データを生成することができる。

【 0 0 2 5 】

図 5 は、本発明による第 2 の実施形態のシステム構成図である。図 4 と異なっており、クライアントには文書専用処理エンジンが搭載されている。図 4 では、クライアントは、復号化して XML 形式のプレーンテキストデータを生成していたが、このようなテキストデータは、その後、処理する際に、再びパーズング処理を行う必要がある。一方、図 5 では、符号化データは、既にパーズング処理が行われたものであるため、クライアントは、文書専用処理エンジンによって直接的に

処理することが可能となる。

【 0 0 2 6 】

図 6 は、本発明による第 3 の実施形態のシステム構成図である。図 4 と異なつて、サーバの XML 形式の符号化変換テーブルは、他のサーバに備えられており、エンコードサーバは、当該他のサーバにネットワークを介してアクセスすることができる。

【 0 0 2 7 】

図 7 は、図 1 の構成について、図 4 のインデックステーブルを用いて複数の文書データを符号する符号化方法の構成図である。

【 0 0 2 8 】

図 8 は、図 4 の構成について、図 3 のインデックステーブルを用いて複数の文書データを符号化するシステムの構成図である。

【 0 0 2 9 】

図 9 は、図 6 の構成について、図 3 のインデックステーブルを用いて複数の文書データを符号化するシステムの構成図である。

【 0 0 3 0 】

前述した本発明の文書データ符号化方法、システム及びプログラムを記録した記録媒体の種々の実施形態は、本発明の技術思想及び見地の範囲の種々の変更、修正及び省略が、当業者によれば容易に行うことができる。前述の説明はあくまで例であって、何ら制約しようとするものではない。本発明は、特許請求の範囲及びその均等物として限定するものにのみ制約される。

【 0 0 3 1 】

【発明の効果】

以上、詳細に説明したように、本発明によれば、拡張可能なテキスト形式の記述言語についての符号化を可能とする。前述したような符号化は、データ伝送量を削減することができるので、特に無線のような伝送速度が比較的遅い通信媒体を介した通信システムに効果がある。また、本発明によれば、拡張可能なテキスト形式で記述された、あらゆる文書データに対して、符号化装置を変更することなく、符号化変換テーブルを置き換えるだけで、それぞれの文書データに適し

た符号化を行うことが可能となる。

【 0 0 3 2 】

また、本発明によれば、復号側装置に文書専用処理エンジンを搭載することにより、受信した符号化データから元の文書データを復元させる必要がなく、復号側装置にとって処理負荷が小さいという効果がある。

【図面の簡単な説明】

【図 1】

本発明の基本的な符号化方法の構成図である。

【図 2】

XML 文書と、本発明による符号化変換テーブルと、該テーブルを用いて該文書を符号化した符号化データとの説明図である。

【図 3】

複数の符号化変換テーブルを結合したインデックステーブルの説明図である。

【図 4】

本発明による第 1 の実施形態のシステム構成図である。

【図 5】

本発明による第 2 の実施形態のシステム構成図である。

【図 6】

本発明による第 3 の実施形態のシステム構成図である。

【図 7】

図 1 の構成について、図 3 のインデックステーブルを用いて複数の文書データを符号化する符号化方法の構成図である。

【図 8】

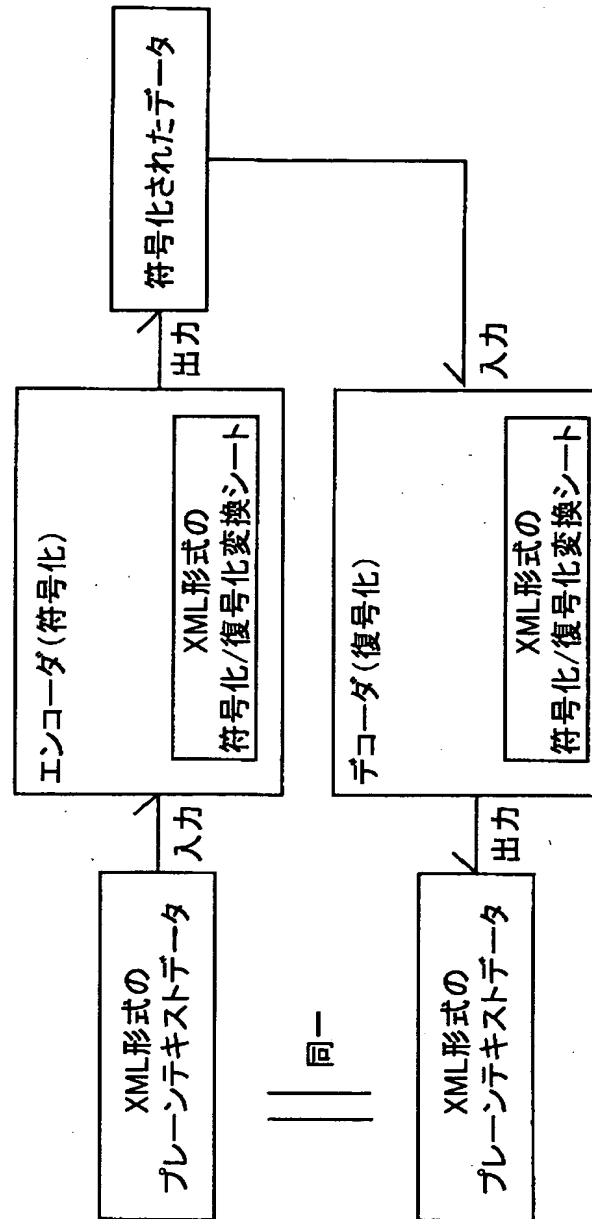
図 4 の構成について、図 3 のインデックステーブルを用いて複数の文書データを符号化するシステムの構成図である。

【図 9】

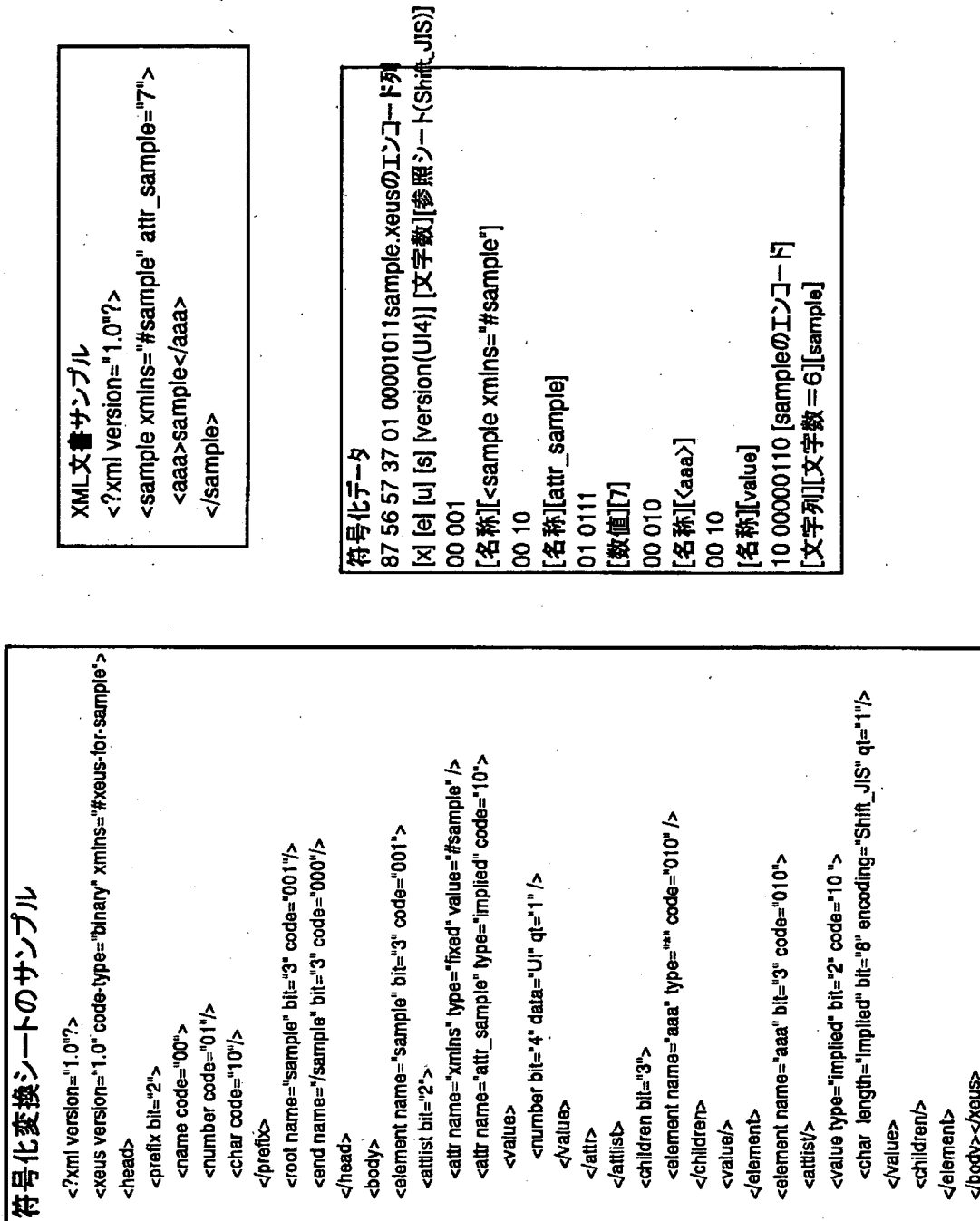
図 6 の構成について、図 3 のインデックステーブルを用いて複数の文書データを符号化するシステムの構成図である。

【書類名】 図面

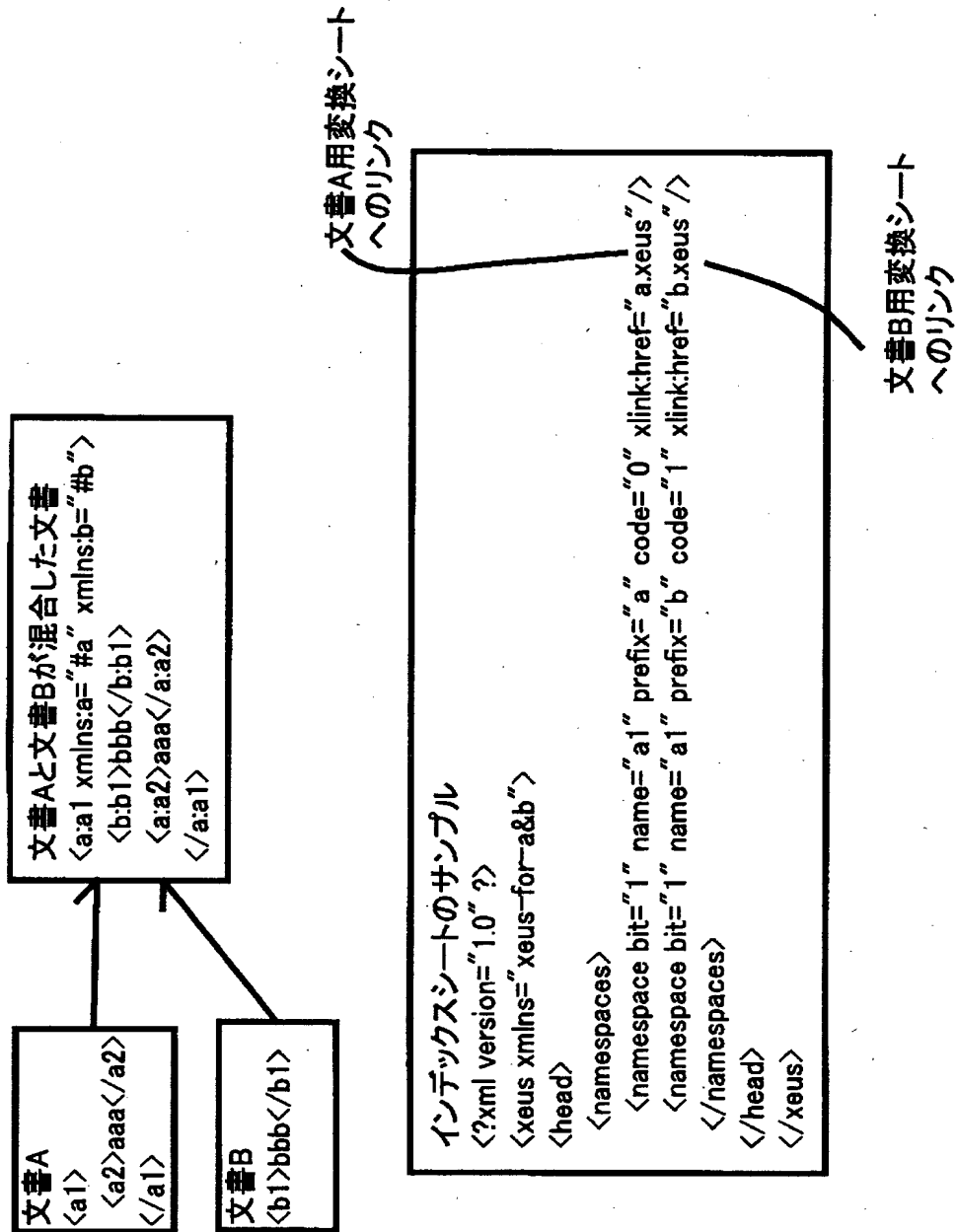
【図 1】



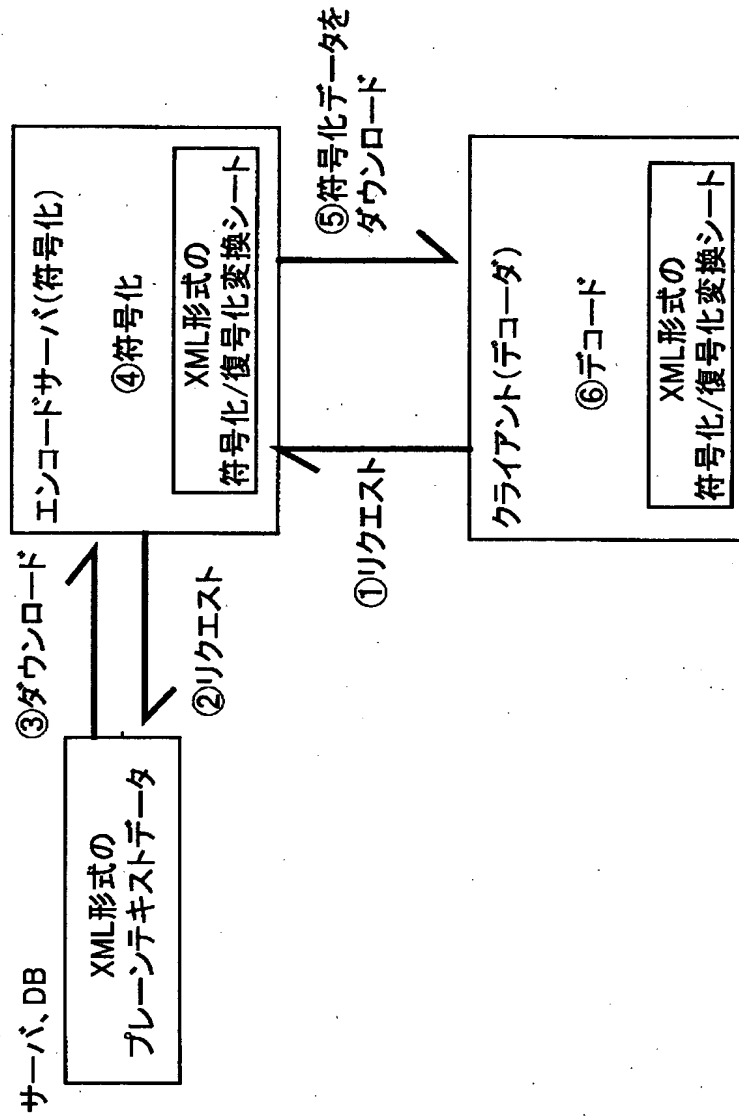
【図 2】



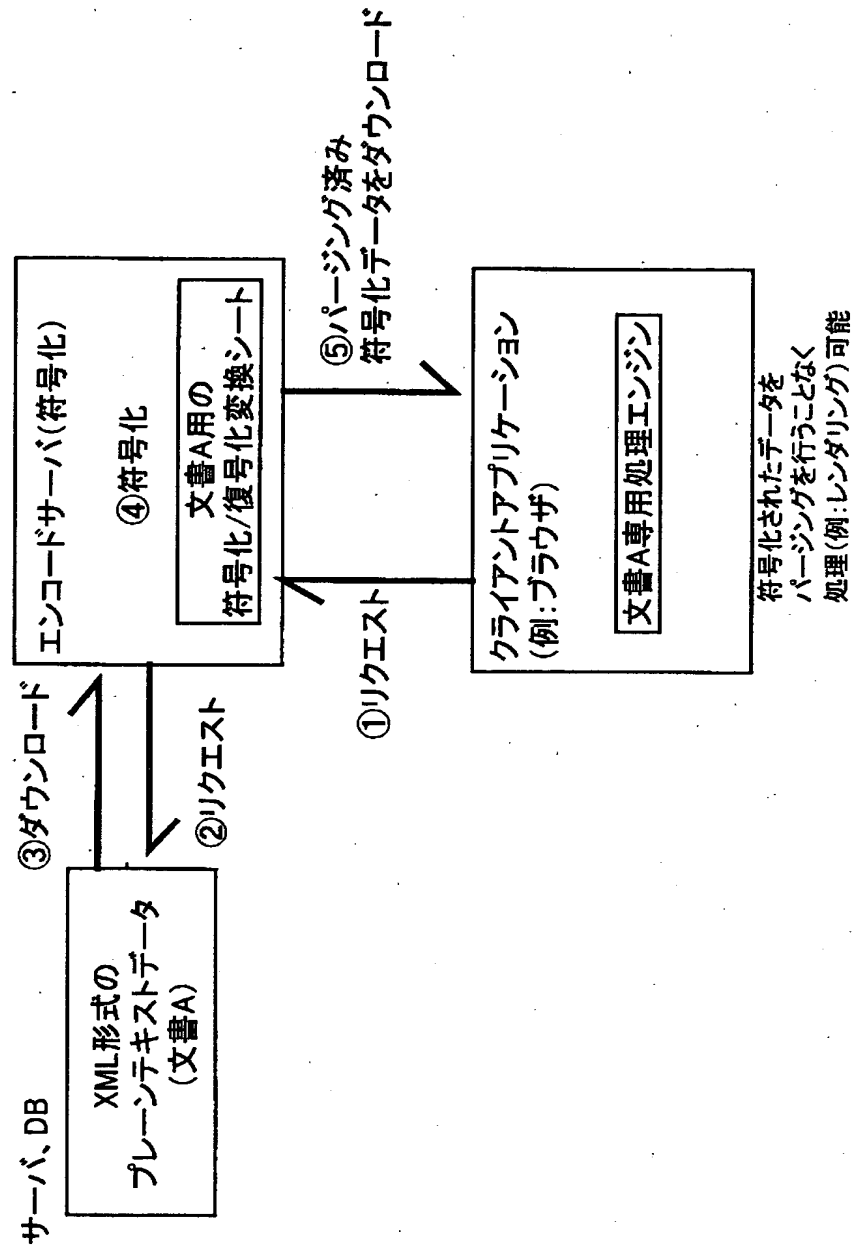
【図 3】



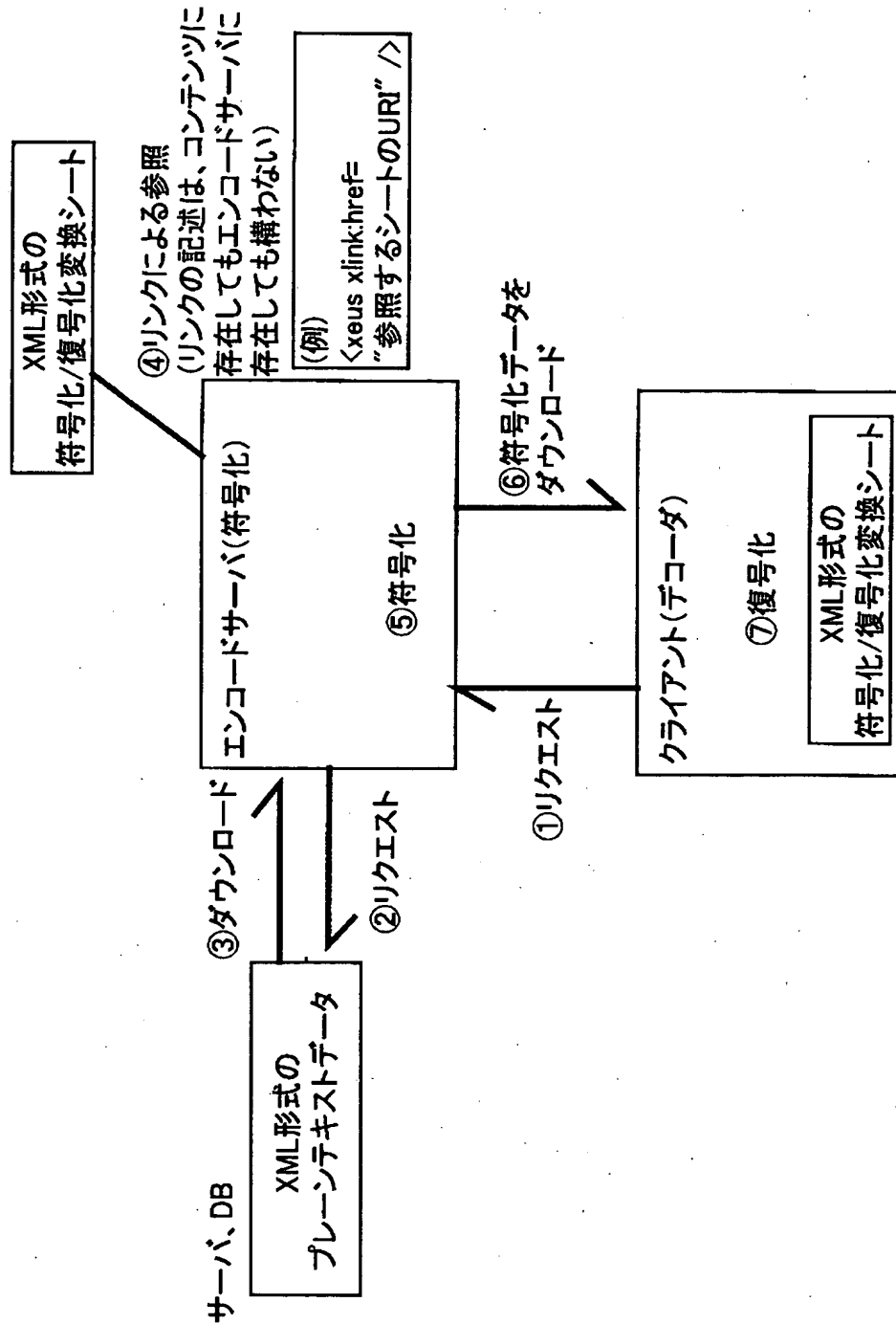
【図 4】



【図 5】

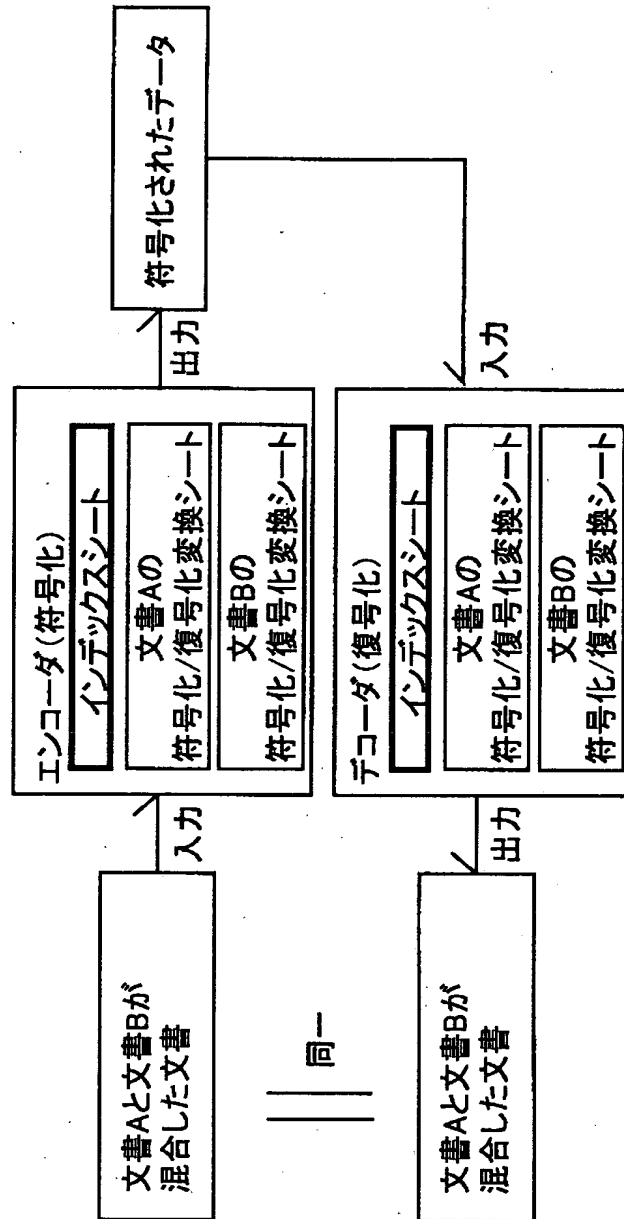


【図 6】



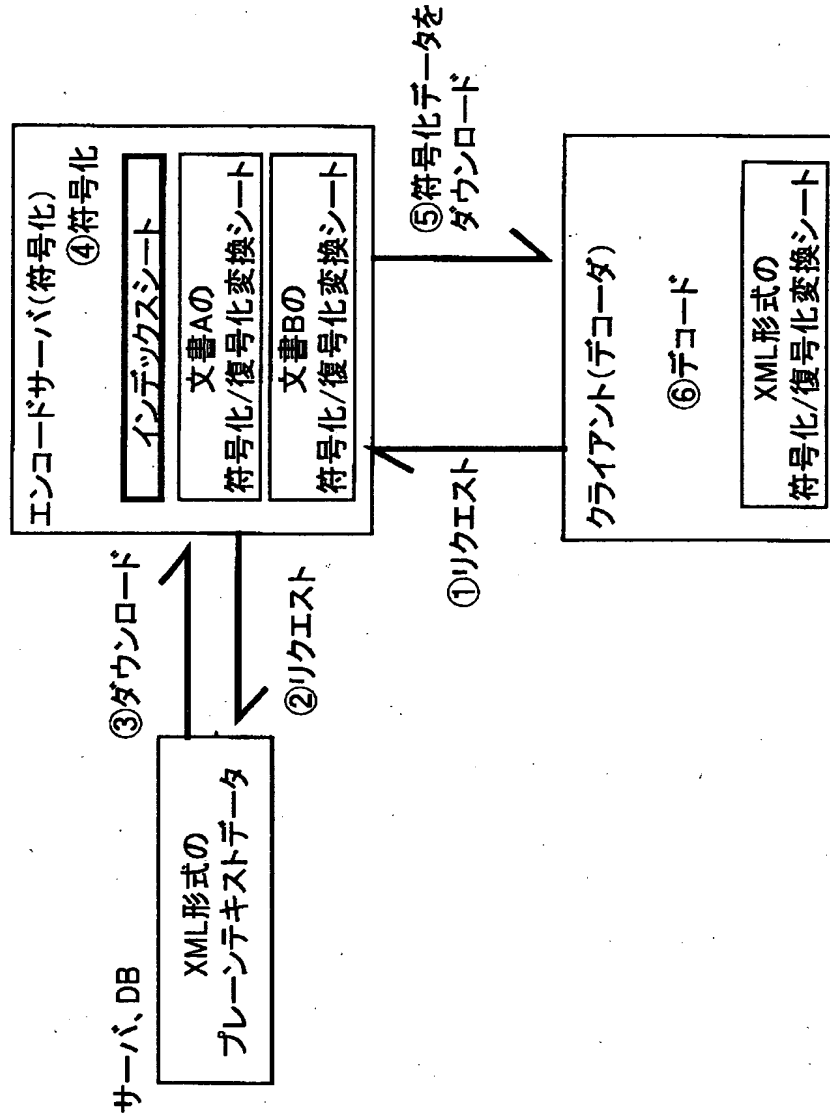
【図 7】

エンコーダ/デコーダに登録された複数の変換テーブル、インデックスシートを参照して符号化/復号化



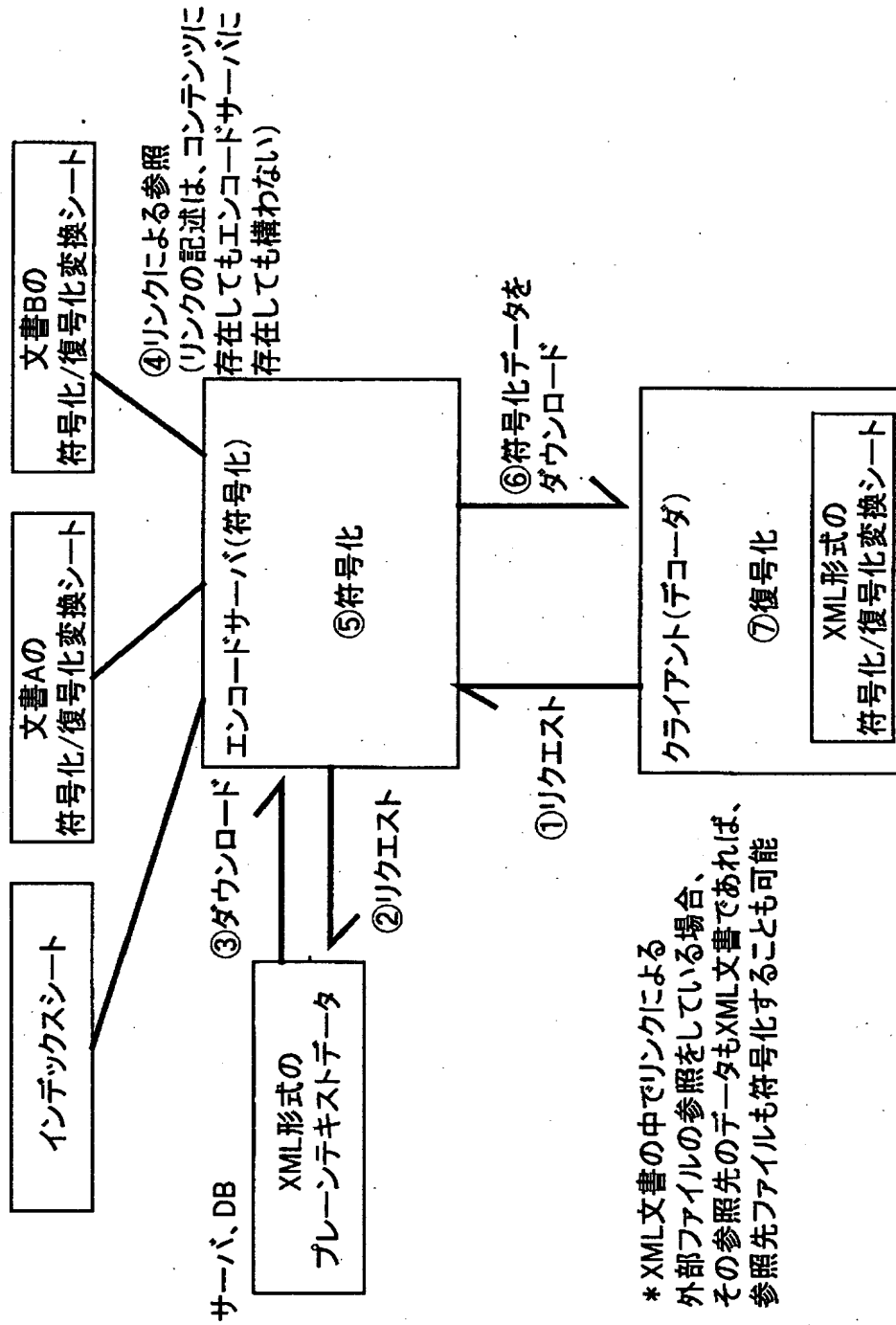
【図 8】

デコーダ(クライアント)、エンコーダ(サーバ)、データがネットワーク上に存在しても構わない



【図9】

変換シート、インデックスシートがネットワーク上に分散して配信されても構わない。



【書類名】 要約書

【要約】

【課題】 拡張可能なテキスト形式の記述言語についての符号化を可能とする方法を提供する。

【解決手段】 拡張可能なテキスト形式の記述言語の文書データを符号化した符号化データを送信し、受信した該符号化データを復号化して文書データを復元する文書データ符号化方法であって、文書データの符号化及び復号化は、符号化変換テーブルに基づいて行われ、該符号化変換テーブルは、名称、数値及び文字列を識別する接頭辞の符号長及び符号と、要素及び該要素に付随する属性の名称、数値及び文字列に割り当てられた符号長及び符号と、要素の親子構造関係とが定義されている方法である。

【選択図】 図 1

出 願 人 履 歴 情 報

識別番号 [000208891]

1. 変更年月日 2000年10月 5日
[変更理由] 名称変更
住 所 東京都千代田区一番町8番地
氏 名 株式会社ディーディーアイ
2. 変更年月日 2001年 4月 2日
[変更理由] 名称変更
住 所 東京都新宿区西新宿二丁目3番2号
氏 名 ケイディーディーアイ株式会社